FIGURE 1

```
.TTTAATCATG GAATATTTCI.AACATACAGA AAAATCACAG AAAATAAATA ACAACCACTC ATTTATCTTC    -1101
TCCCCAACCC CATGTAATAA ATATTAAAAT ATTGTGTTAA ATGCTAAATT TAACACATGC TAAAGGTTCC     -1031
TGGCTGGATG TGGTGGCTCA CGCCTGTAAT CCCAGTACTT TGGGAGGAGG AGGTGGGAGG ATTGCTTGAG      -961
TCCAGGAGCT CGAGACCAGC ATGGGCAACA TAGTGCGATC TCGTCTCTAC AAAAAACAAA AAAATTAGCT      -891
GGGCATGGTG GTGTGCATCA GTAATCCCAG TGACTGGGAG GCTGAGGTGG GAGAATTGCT TGAGTCTGGG      -821
AATTTGAGGC TGCAGTGAGC CCTGATCATG CCACTGCATT CCAGCATGGG CGACATAGCA AAACTTGTCA      -751
AAAAAAAAAA AAGTTTCCTC TCTGCCCCAC CATAGACAAC CACTCTTCTG ATTTCTATCT TCGTAGATGA     .-681
                                     AP1
ATTTTGCCCA TTCTCTTGTA TATGAAAGGA ACCAGACATT AGGCATTCTG GTGTCTGGTT TCTTTCACTT      -611
AAGATAAAAT TGAGTTAACC TGTATTGTTG TACAGAACTG CAGTTTGTTC TTTGTTATTT ATTGTAAAGA      -541
CAGGGTCTGG CTATGTTGCC TAGGCTGGTC TCGAACTGTT GGCCTCAAGC AATCCACCTG CCAAGCTCTG      -471
GGACCACAGG CATGAGCCAT GGCATCTGAT CKGTAGTTTG ATCTTATTTC TTGCTGAGTA GTAGCCCATG      -401
   AP1
GCATGACTTT ATTATTTTGG GTGTCCATTC TCCTCTGGAG GGGCTCTGCT TTTTGAAACC ACACCCTGGC      -331
       Ets
CTAGCTCCCC TTCTCCCTGC CTCTCTGCAG GCTCACATCC ACATGCCAAG ACCTCTGCAG CCATTCTGCT      -261
       Ets
TCCTGTCCTT CCACTCCTGT GGGACCTCAG AGAGCTACGG GGCTCCCTGG GTACCAACTG GCTCCTGAGG      -191
   Sp1/Sp3                                              Sp1/Sp3
CCTGGGGGAG GGTGGTCTTC TGGGAGAAGG AAGCCAGGTC CCTGCAGGTT GTGGAGGGGG ACAGAATGAG      -121
           Sp1/Sp3                                                ETS
GGTTTTTCCC CAGGATGTTG TTGGCCCCTG CCCCCACTTC TGTTCCATAA TTAACCACGC CCCTCCTACC       -51
           Sp1/Sp3                      +1
CACTGTGCCC CTCTTCCTGC TGTGTGGAGG CCCTGAATCA TTATTTTAAC TACCCCCTGG GAGGGTGAGC       20
   Ets              Ets
ACCTTCTGTG CTCTGTCCCC AACCTTCCAC TTCCCCTCAA CGCGCTGCTC AGGG|ATG|ACC TTCGGCACTG       90
                                                          M  T   F  G  T
TGCTTCTTCT GAGTGgtaag tggggccagg gtgctgggga gaagcttgga ggagttctga ggggactcca      160
V L L L   S
tctgggaggg caggctgggg gctggtggtc ggctccaacc actcttatga ggagctgagg caggggagtg      230
cttcatgtgc gagtggcccg gagtcagtag agtgtgacct gaatgaagag gggctcaggg gctgtgctca      300
ggtggcgact aagctacctc tccagctggc tatgttgtcc caggcttccc tgctcccact catggagtcc      370
ctggtgtggg tgacagaggt ctccccagcc tcccccggga gtggaaggcc acagaagcca ccagggaggg      440
ggaaaggttg gacatcacct ccctgggcct nnnnnttccc ccaagtcctg actgcacgta gggaagaggc      510
                                  INTRON 1
ccctgctga aaactgcatc agagtcacat tcacgtgcca tcaaaaatca ggcttggctg ggtgcggtgg       580
ctcatgctta taatcccagc actttgggag gccgagatgg gcgtatcccc tgaggtcagg agtttgtgac      650
cagcctggcc aacatggtga aaccccatct ttaccaaaaa tataaaaatt agccgggcat ggtggcgtgc      720
acttgtaatc ccagctactt gggaagctga ggcaagagaa tcgcttgaac ccaggagacg gaagttgcag      790
tgagctgaga tcgtgccgtt gcactccagc ctcagcaaca gagcgagact ccatctcaaa aaaaaaaaaa      860
aaaaaagaaa aaaaagaaa aagaggctgg gaggtcctag ggattggggc ttctttaact cccagcctcc      930
ccgcccacca aatattcctc agTCCTGGCT TCTTATCATG GATTCAACCT GGATGTGGAG GAGCCTACGA     1000
                      V  L  A   S  Y  H  G  F  N  L   D  V  E   E  P  T
TCTTCCAGGA GGATGCAGGC GGCTTTGGGC AGAGCGTGGT GCAGTTCGGT GGATCTCGgt aggccccact     1070
I F Q E   D  A  G   G  F  G . Q  S  V  V   Q  F  G   G  S  R
------------------------INTRON 2 (3019 bp)------------------------
cccccaagtg cccgctgctc ccacccctcc tgtggctgca gtgacatggc catggttgtg tctccagACT     4080
                                                                          L
CGTGGTGGGA GCACCCTGG AGGTGGTGGC GGCCAACCAG ACGGGACGGC TGTATGACTG CGCAGCTGCC     4150
  V  V  G . A  P  L   E  V  V  A    A  N  Q    T  G  R   L  Y  D  C   A  A  A
ACCGGCATGT GCCAGCCCAT CCCGCTGCAC Agtgagtgac cacctgggaa ttgggcccct caacctcct     4220
  T  G  M   C  Q  PI    P  L  H       INTRON 3
ggacccaact gtgccccgc ttagcttcca gtccagacct tccccgcaaa tgagtgtgtg ctgtgagtga     4290
gacccgcgt gtctgccctt gcagTCCGCC CTGAGGCCGT GAACATGTCC TTGGGCCTGA CCCTGGCAGC     4360
             I  R   P  E  A  V   N  M  S   L  G  L   T  L  A  A
CTCCACCAAC GGCTCCCGGC TCCTGgtgag tgagtgtctt gggccacggg ggggtggggt ggggcgggg     4430
  S  T  N   G  S  R   L  L       INTRON 4
gtgttgttgg ggaggaggct ggggctggga gtgaaggagg aggggctgct agggactcct ggctcacagg     4500

cttctgcctc cagGCCTGTG GCCCGACCCT GCACAGAGTC TGTGGGGAGA ACTCATACTC AAAGGGTTCC     4570
             A  C   G  P  T  L   H  R . V   C  G  E   N  S  Y  S    K . G  S
TGCCTCCTGC TGGGCTCGCG CTGGGAGATC ATCCAGACAG TCCCCGACGC CACGCCAGgt aggtccctgg     4640
  C  L  L   L  G  S  R   W  E  I    I  Q  T   V  P  DA     T  P .
--------------------------INTRON 5 (4267 bp)--------------------------
caggagctgc aggaggggt tgggcccccg cagtgcatct ccgattcctc cccattcccc cacacAGTGT     8840
                                                                        E  C
CCACATCAAG AGATGGACAT CGTCTTCCTG ATTGACGGCT CTGGAAGCAT TGACCAAAAT GACTTTAACC     8910
  P  H  Q    E  M  D  I   V  F  L    I  D  G    S  G  SI    D  Q  N    D  F  N
AGATGAAGGG CTTTGTCCAA GCTGTCATGG GCCAGTTTGA GGGCACTGAC ACCCTGgtga agactgggca     8980
Q  M  K  G   F  V  Q    A  V  M    G  Q  F  E    G  T  D    T  L
--------------------------INTRON 6 (1255 bp)--------------------------
aacaatagta acaggcactg agccctgggc cctccccact ggcctttgca gTTTGCACTG ATGCAGTACT    10240
                                                          F  A  L    M  Q  Y
CAAACCTCCT GAAGATCCAC TTCACCTTCA CCCAATTCCG GACCAGCCCG AGCCAGCAGA GCCTGGTGGA    10310
  S  N  LL     K  I . H   F  T  F    T  Q  FR     T  S  P   S  Q  Q    S  L  V  D
TCCCATCGTC CAACTGAAAG GCCTGACGTT CACGGCCACG GGCATCCTGA CAGTGGTgta aagcaacccc    10380
   P  I  V   Q  L  K   G  L  TF     T  A  T   G  I  L   T  V
gacccca-----INTRON 7
```

FIGURE 2

FIGURE 3

+1

CCTGAATCATTATTTTAACTACCCC

mRNA ————————————————————————————————— ATG ————————▶

←— N470

———————————————————————————— 93 nt

———————————————————————————— 92 nt

←— N379

———————————————————————

——————————————————
———————————————
————————————— 75-79 nt
————————————

———————— 51 nt

—————————————
————————— 71-74 nt
———————
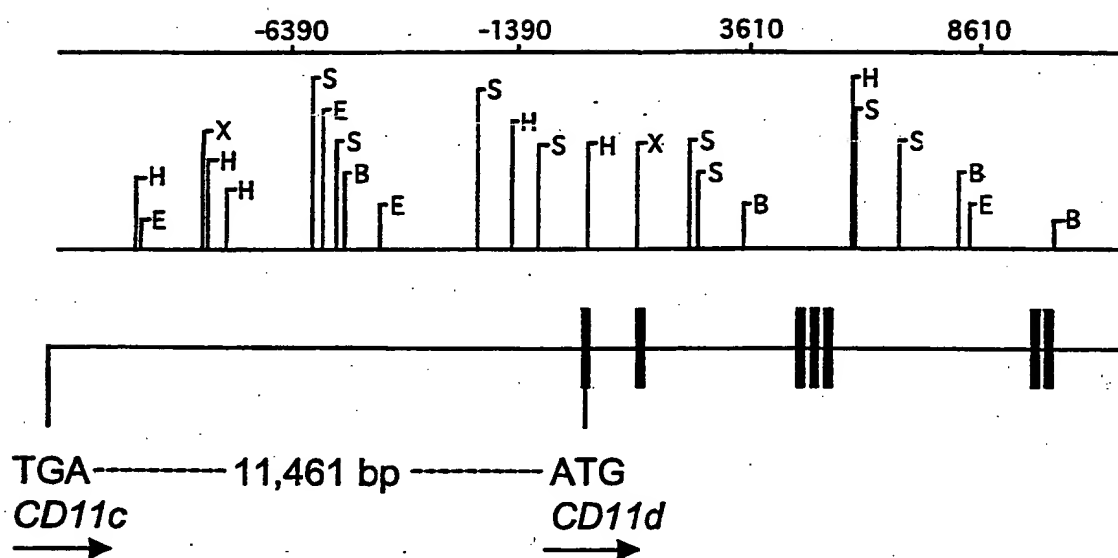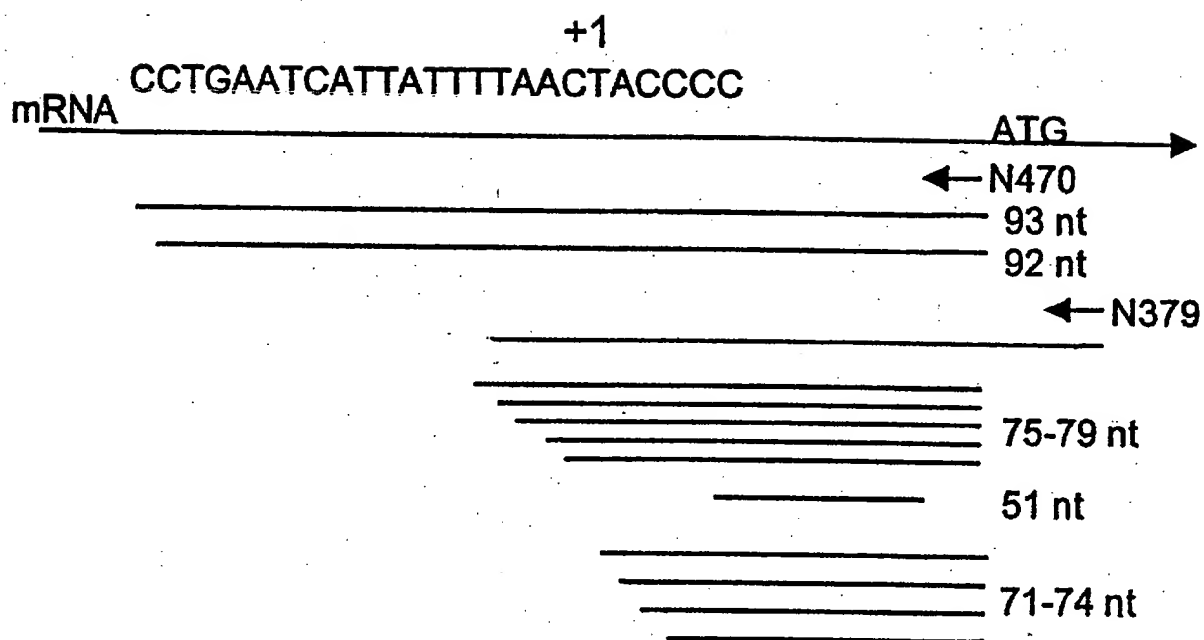
FIGURE 4

FIGURE 5

FIGURE 6

% Luciferase Activity

FIGURE 7

FIGURE 8

1   2   3   4   5   6   7   8   9   10

```
-76              -66                    -54
 |                |                      |
CCATAATTAACCACGCCCCTCCT
              ●○  ○  ○○○○  ○○
-53                     -38         -31
 |                       |           |
ACCCACTGTGCCCCTCTTCCTGC
   ○  ○           ○●  ○  ○
```
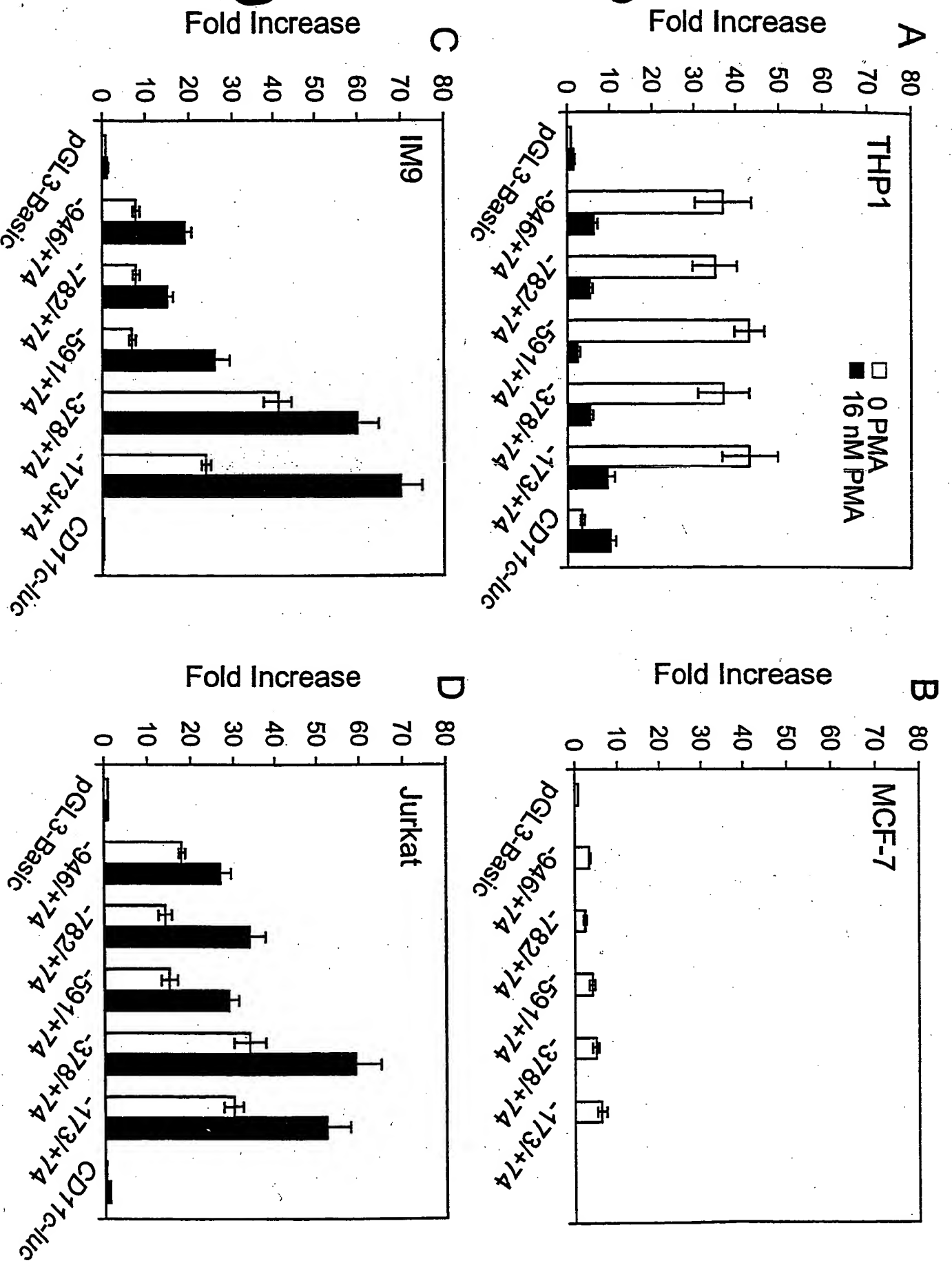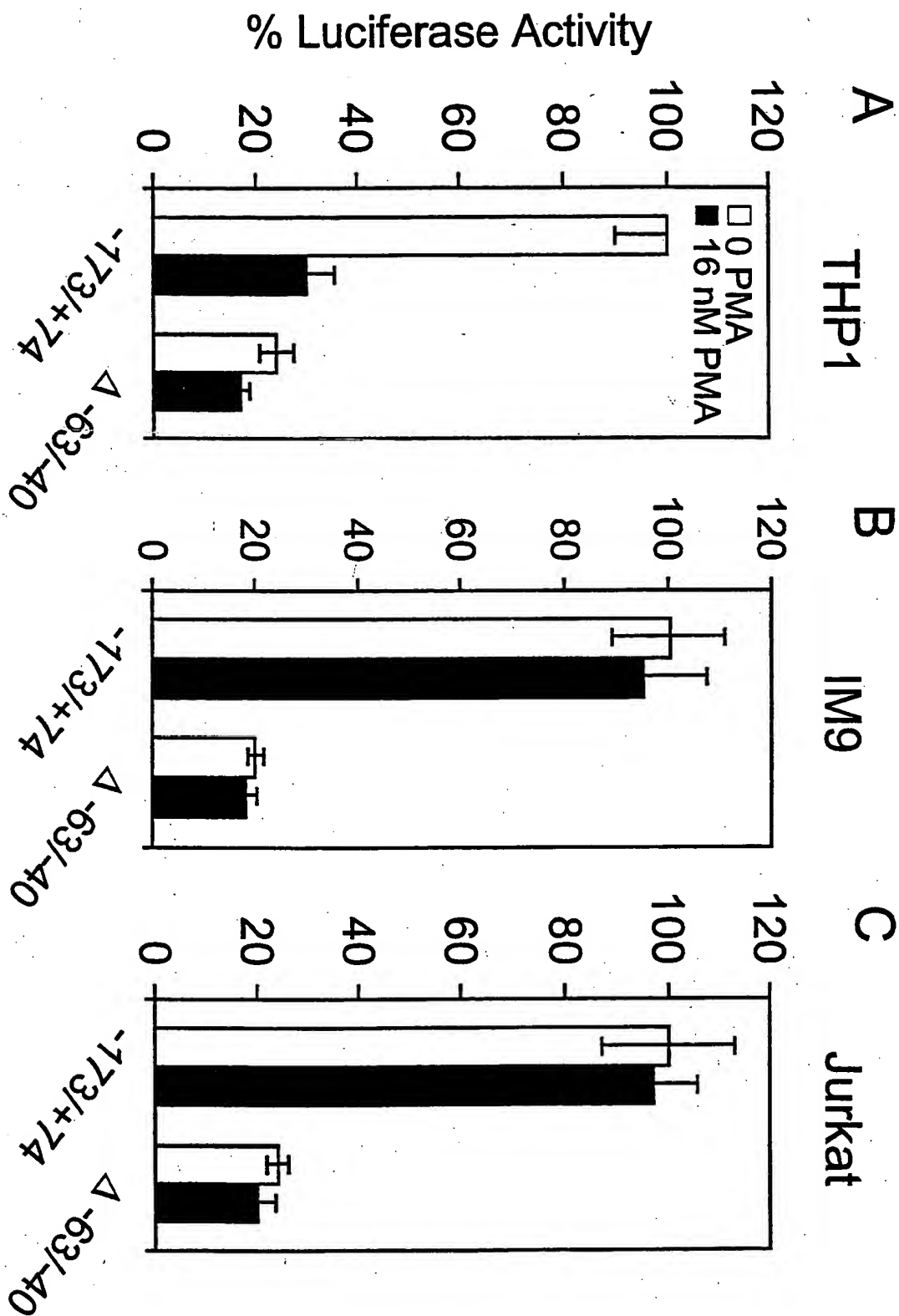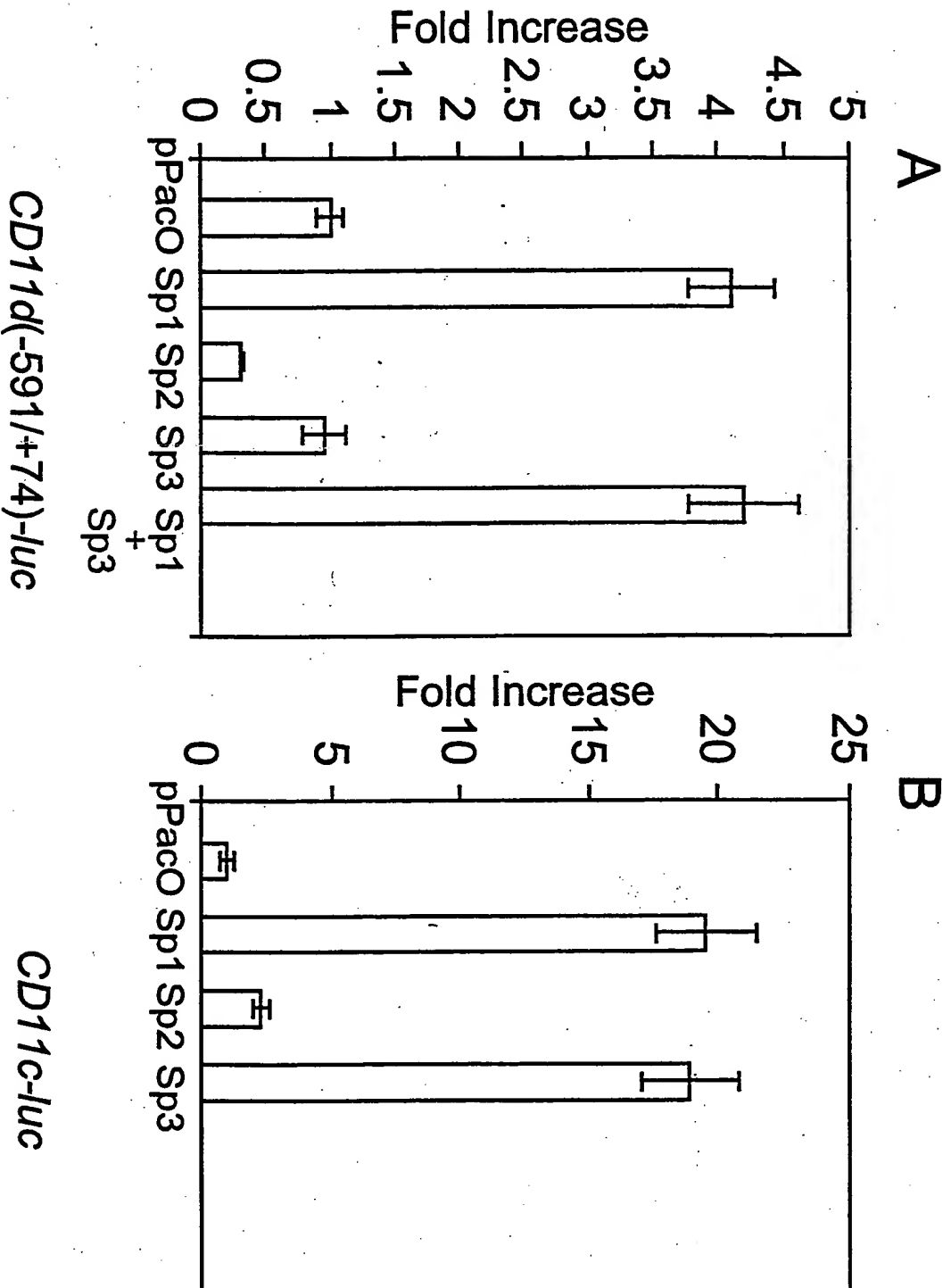
FIGURE 9

TRAnslATioNAl
Stop
Codon
FoR
CD11c

                                                                    -11321
<u>TGA</u>TCCCTCT TTGCCTTGGA CTTCTTCTCC CGCGATTTTC CCCACTTACT TACCCTCACC TGTCAGGCTG

                                                                    -11251
ACGGGGAGGA ACCACTGCAC CACCGAGAGA GGCTGGGATG GGCCTGCTTC CTGTCTTTGG GAGAAAACGT

                                                                    -11181
CTTGCTTGGG AAGGGGCCTT TGTCTTGTCA AGGTTCCAAC TGGAAACCCT TAGGACAGGG TCCCTGCTGT

                                                                    -11111
GTTCCCCAAA AGGACTTGAC TTCGAATTTC TACCTAGAAA TACATGGACA ATACCCCCAG GCCTCAGTCT

                                                                    -11041
CCCTTCTCCC ATGAGGCACG AATGATCTTT CTTTCCTTTC CTTTTTTTTT TTTTTCTTTT CTTTTTTTTT

                                                                    -10971
TTTTTTGAGA CGGAGTCTCG CTCTGTCACC CAGGCTGGAG TGCAATGGCG TGATCTCGGC TCGCTGCAAC

                                                                    -10901
CTCCGCCTCC CGGGTTCAAG TAATTCTGCT GTCTCAGCCT CCTGCGTAGC TGGGACTACA GGCACACGCC

                                                                    -10831
ACCTCGCCCG GCCCGATCTT TCTAAAATAC AGTTCTGAAT ATGCTGCTCA TCCCCACCTG TCTTCAACAG

                                                                    -10761
CTCCCCATTA CCCTCAGGAC AATGTCTGAA CTCTCCAGCT TCGCGTGAGA AGTCCCCTTC CATCCCAGAG

                                                                    -10691
GGTGGGCTTC AGGGCGCACA GCATGAGAGC CTCTGTGCCC CCATCACCCT CGTTTCCAGT GAATTAGTGT

                                                                    -10621
CATGTCAGCA TCAGCTCAGG GCTTCATCGT GGGGCTCTCA GTTCCGATTC CCCAGGCTGA ATTGGGAGTG

                                                                    -10551
AGATGCCTGC ATGCTGGGTT CTGCACAGCT GGCCTCCCGC GGTTGGGTCA ACATTGCTGG CCTGGAAGGG

                                                                    -10481
AGGAGCGCCC TCTAGGGAGG GACATGGCCC CGGTGCGGCT GCAGCTCACC AGCCCCAGGG GCAGAAGAGA

                                                                    -10411
CCCAACCACT TCCTATTTTT TGAGGCTATG AATATAGTAC CTGAAAAAAT GCCAAGCACT AGATTATTTT

                                                                    -10341
TTTAAAAAGC GTACTTTAAA TGTTTGTGTT AATACACATT AAAACATGCA CAAAAAGATG CATCTACCGC

                                                                    -10271
TCTTGGGAAA TATGTCAAAG GGTCTAAAAA TAAAAAAGCC TTCTGTGGAT ATGAGTCCTG AAGGATGACA

                                                                    -10201
CCCATGGGGT CCCTTTACCA CGGTGGACCC TGGCCAGCAC TGAGGCCTGG GGCCAGGACA AGAAGTTAAC

                                                                    -10131
CAGAGTAGGG TTGTGAATAT CCCTCTCTTG GAAGTAACCT GACCTCTTAA TCTGCTCACT CCACTCTCAG

                                                                    -10061

GGCTGGTGCC GATGGTAAGC TGGTGGAGCT GTCGGGTGGA GGGGGCATAG AATAGAGAAG GGACAACCTC

-9991

CAGTGGCTAC TTTTCCACCT GGAAAGGTCT CTGGAGTGAC CAATACTCAC AAGCGTTTCC TACAAGTCCT

-9921

AGGATGTGTT GAAGGGCACA CTGTCTGCAT ATAGTGAGTG ATTGAAGAAC ATGTTGGGGT CCCACATTGA

-9851

GAGCTGCTGC CCACAATAAG GTCATTCTTG CTATTATGCC ACCATCCTGG CATAAAGTTC ATCATGGTGC

-9781

TTGGCACTGA GCTGGGGGCC TCACAGGACA AGCCATTCCT GACCTCGGAG TGACGCCACT GCAGCTATCA

-9711

CCAGCAAGCG ACCCGGGCCG TGTGGATGTT TCAATTAGAA AAACAGAAGG GAGGCAGTTG AGTGATTTGA

-9641

AGGGAAGATG GAAAGTGGCC CTTTACCTCC AGCCAAAAAT GTCTGTCCTA TACATCAGCA GAGGCTCCAA

-9571

AATCCCTGTG GATTTTGAAG CTTTTGAGTC CCCAGGATGA CTAATTATTA TGCAGTTTCC TCAGAAAGGG

-9501

AATCAGAAGA TAAGGCTTTG TAAGAATTCA GCCCTAATGG CTGGGCACAG TGGCTCATGC CTGTAATCCC

-9431

AGCACTTTGG GAGGCCGAGG CAGGAGGATT GTTTGTGCTC AGAAATTTGA GACCACCCTG GGTAATATAT

-9361

TGAAACCTTG TGTCTACAAA AAAATTTAAA AATTACCCAG GCATGGTGGC ATGTGCTTGT AGTCCCAGCT

-9291

ACTTGGTAGG CTGAAGCAGG AGGATCACTT GAGCCTGGGA GGTTGAGGAT ACAGTGAGCT GTGATTTGGA

-9221

CCACCACACT CCAGCCTGGG CAACAGAGAA AGATCATGTC TCAGAAAAAA AAAAAAAAAT TGACCCTAGA

-9151

GTGGTGTTTC TCAAAATGTG TTCCACGAAC CACTGGTGGT CAATGATGGT CTTCTAAGTG GAAGGTTTTA

-9081

GAGAAAAAGA GCAAGAAACC CATACATCTC AAACATTTGA AACTAGTGAT TTGCACAGAA ATAGTGTTGT

-9011

GGCCTTAATA ATTGTGTGGC ACACGGACTC CAGGGACTAC AGTGGGTTCT TGTCTAAATT CAGGCAACAA

-8941

GTTGTTATTT TCTATTTTAT TTTATTATTA TTATTTTTTG AGATAGTCTC ACTTTGTCTC CCGGGCTGGA

-8871

GTGTAGTGGC ACGATCTCGG CTCAACGCAA CCTCTGTCTC CTGGGTTCAA GTGATGCCTC TGCCTCAGCC

-8801

TCCCAAGTAG CTGGGAGTAC AGGGGCGTAC CACCATGCCC ATTTTTATTT ATTTATTTTT GAGACAGAGT

-8731

CTCGCTCTGT CACCCAGGCT GGAGTGCAGT GGCATGATCT TGGCTCACTG CAACCTCCGC CTCCCAGGTT

-8661

CAAGTTCAAG CGATTCTCCT GCCTCAGCCT CTGGAGTAGC TGGGATTACA GGCAGGCACC ACCATTTCCA

-8591

GCTAATTTTT GTATTTTTAG TATAGATGGG GTTTCACCAT GTTGACTAGG CTGGTCTCGA ACTCCTGACC

-8521

TCATGATCCG CCCTCCTCGG CCTCCGAAAG TGCTGGGATT AGAGGTATCA GCCACTGTAC TTGGCCGACA

-8451

AGGTGTTATT TTCTGATATT CTTCCTTTGT GTGTTATTGT GTACATTTGT TACATTTGCA TTTTCAGGGT

-8381

TGGCTATTGT GTTGCATTAG ATCCCCGAAT CACAAAATGG ATCAATGGCT CAAAAGCATG GAAGTTGTGA

-8311

TTAAAAACTA ATCTAATTGC TACAATTTAC AATAATGTCA TCAAAGTCAA TATTGACTTT TAAATATTGA

-8241

GCCCAGTGCA CGTATAGTAT AGACATGCAT ACCGGAATAA GTGATTGTGA GCCAAAACCC GAAAATATCT

-8171

AGAAGGTATT ATACTCCCTG ACAGGTAGGT TGTATTGGTT CTGACATGTA TTTGTCCCTA GTGTGCTGCC

-8101

CATTCTGAAA CTTTATCAAA CAGTCGCATG AACCTCTGAA AGCTTTTGTG TTATTTTCTT ATTTATTTAT

-8031

TTATTGAGAT GGAGTCTTGC TCTGTCGCCC AGGCTGGAGT GCAGTGGCAT CATCTTGGCT CACTGCAACC

-7961

TTTGCCTCCT GGGTTCAAGT GATTCTCCTG CCTCAGCCTC TTGAGTAGCT GGGATTACAG GCGCGCACCA

-7891

CCACGCCCAG CTAATTTTTG TATTTTTAGT ATAGACGGGG GTTTCACCAT GTTGGTCAGG CTGGTCTCGA

-7821

ACCCCTGACC TCATGATCTG CCTGCCTCAG GTAAAGCAAT AGAGATTCTT AGAACAACTG CTACATGTAG

-7751

CTTTCCTATT CAAAAGTGAT TAGTGTTGTC ACCGAATACA GAGGAGACAG CAAAACCACA GTGACATAAA

ESN490   -7681)

TCAAAGGTGC TTTTTAAAGT AGCAAAAGTA GGTACAAGTC ACATAATTTC CAAGAAGCTT GTAGAAATGG

-7701

-7611

CAGTAGAGTT CATACCTGCT ATTGAAAGGT TGCTTTTGGC TGCAAATAAT AGAAAAAAAC AAAAGCATGT

-7541

AAGAGCAGAC AGAAGACCTT TACTCTGCAA GAGGTTCAGG TGCAGGTTAG TGTTTAATGC AGAGTCTCAG

-7471

CATTGACAGA TTCTTTCTGA TCTTCCAATT GATCGTCCTT GCGGGGGCGG TTTAGTTCTT TCCCACTGAC

-7401

TAGGATTGGG TCAAATTCCA TCCCCTTGGT TGCATGCAGT GCTGAGAAGG TGAGCATGTG CTTTTCACAG

-7331
GCTTAATAAA AAGAGGTAGC TCCAGCCAGG TGCAGTGACT CATGCCTATA ATCTCAGCAC TTTGGGAGGC

-7261
AGAGGTGGGT AGGTCACCTG AGGTCAGGAG ATTGAGAACC AGCCTGACCA ACATGGCAAA ACTCTGTCTC

-7191
TACCGAAAAT ACAAAAATTA GCTGGGCATG GTGGCAGGTG CCTGTAATCC CAGCTACTTG GGAGGCTGAG

-7121
GCAGGAGAAT CGCTTGAACC TGAGAGGTGG AAGTTACAGT GAGCTGAGGT CATGCCACTT GCACTCCAGC

-7051
CTGGGGGACA GAGTAGAACT CTGTCTCAAC AAAAAAAAAA AAAAAAAGAG AAAAAAAAAG GAGGGTAGCT

-6981
CCACCAGCCA GGAAGGTGGC AGCGCTGGTG GCTGTTGGAT AGGCTACCTA CAGTGTCTGG CAAATACTAT

-6911
GCTTGAAGAC TATGCTGTGA GCAAGATTCC TTTGTGAAGG AACAGCTTGG ACATTGTGTA TGTCAGAGGT

-6841
ATACAGCAGA ATAGCAGTGA CTAACGCTTG TGTGGGAGAG CAAGCATGTC ACCTCATACT TGGAATAACT

-6771
CACTGCCATA CAAAGTCTGA ATCAGCTTTC GTCTTTGTGC AACACATGTA TGTGGGAGCT TTTCAGCTGC

-6701
TGAAACCTCT AGTGACAGAA AAGGAGGTTT TGTTGTTCAT TTGTAATTAA TGTTAATCCT ATGAGTGGTG

-6631
GGAGAGATAG TGAGGTAGGA GATCAGCAGG ACCTGTTTTC TGGTCACAAC CCAGCTAATC AGAGCATGAT

-6561
CTGGTCAAGA TGGGATGCAC TAAAAAAACA GCCCAAACCA GCAGATGGCC AGGAAAGCAA ACTCTCATTA

-6491
CCCTCGCCAC TTATTAGCAT AAAGACACTC CCACCGGTGC CATGACAGTT TACAAATGCC ATGGAAACAC

-6421
ACCATAGCAA CGGTCAGCAA GTTACCTCAT ATGGTTCTGG AAACTCCCCA CACCTTTTCC AGATAGTTCT

-6351
GAATAACCCA CCCCTTAATT TGCATGTAAT TAAAAGTCGG TATAAGTACA GTTAGCCAGC AGCCCACTGG

-6281
CTGCTACTGT GGGCTCACTG CCTATGGGTT GTCCTGCTCT GCAAGGAACA GCTACCTTGC TGCCACTGCT

-6211
GCTTCAATAA ACCTGCTTTC TTCCACCACA GGCTCGCTCT TCAGTTCTTT CCTGAGCAAA GTTAAGAACC

-6141
CTCCCGGGCT AAGCCCCAAT TTTGGAGCTT GCCTGCCCTG CATCAGTAGA ATGGGCTAAC TACTTACGGT

-6071

GCACTCAGGC TAAAGAGGCT GATGCTTGCA GGGCAGTATT CACAGAGCAC ACGGTAGTTC ACGGGATGCC

-6001
TCTCACCCTT GACTCAGTGC TTAAGAAAGG AGGGAAAATG GTGAACATGA TCAAATCATG GCCATTGCCT

-5931
ATTCATCTTT TCAGTGTTGT ATGGAGGAAT AGGCAAGTAG GAGATTGCTT TTCACATTAA TGTCAAAGAG

-5861
AAAGATAGTT ACTTGGAACT TAAAAAAATT AATTGTGATA AAATATACAT AACATAAAAT TTACCATCTT

-5791
AACCATTTTT AAGTATAGCC AATCTCAAGA GCTCTTTCTA TCTTGTAAAA CTGAAACCCT ATACCCATTA

-5721
AACAACTCCC AATTCTCCCC TTTCCCTAAC TCCTGGCAAC CACAATTCTT TCTGTCTCTA TGAATTTGAC

-5651
TGCTTTGGCA TGTCATAGAA ATAGACTCAT ACAGCATTTG TCTTTTTGCG ACTGGCATAT TTTGCTTAGC

-5581
ATAATGTCCT CAAGGTTCAC CCATGTGGTA GCATGTGTCA GAATTCCTCT CCTTTTGAAG GCTGAATAAT

-5511
ATTCCATTGT GTGTATATAC CACGTTTTGT TTATCCATTT GCCCATCAAT GGGCATTTGG GTTGCTTTTT

-5441
TTGCCTCTCA TGAATGATGA ATATGGGCGC ACAAATATCT CTTCAAGACC ATGCTTTCAA TTCTCTTGGG

-5371
TATACACCCA GAAGTGGAAT TGCTGAATCA TATGGTAATT TTTTTTTTTT TTTGAGACAG AATCTTGCTC

-5301
TGTTGCCCAG GCTGGAGTGC AGTGGCACAA TCAGAGCTCA CTGCAGCCTT GGTCTTCTGG GCTCAAGCGA

-5231
TCCTCTTGCT TCAGCCTTCC GAGCTTCTGG GACTAAAGGT GTGTGCCATC ATGCCTGGCT AATGTTTTAA

-5161
AAACGTTGCC AGGCATGGTG GCTCGTGCTT GTAATCCTAG CACTTTGGGA AGCTGAGGCA GGTGGATCCC

-5091
CTGAGGTCAG GAGTTTGAGA CCAGCCTTGC CAACATGGTG AAATCCCGCC TGTACTAAAA ATACAAAAAT

-5021
TAGCTGGGTG TGGTGGCATG TGCCTGTAGT TCCAGCTACA GGCAGGAGAA TTGCTGGAAC CTGGGTGGCA

-4951
GAGGCTGCAG TGAGCCGAGA TTGCACCACT GCACTCCAGC CTGAGTGACA GAGTGAGACT CTGTCTCAAA

-4881
AAAAAAAAAA ATTTTAGAGA TGGTGTCTCA CTGTGTTGCC CAGGCTGGTC TTGAACTCCT GCCCTAAAGT

-4811
GATCCTCCTG CTTCCGCCTC CCAAAGTGCT GGGATTACAG GCATTAGCCA CCATGCCTGG CCTAGCTAAA

-4741

TTGTCTTTAA TGTCGCATGT CTGCAAAAAA CACATCTATA AAGCTAGAAA AGTTGAGCAT CCAACTTTTT

ATGATTTAAC TCTCATGACC TGGCAATTTT TCTAGCAAGG AGCCTGGGCT GGTGGTTTTA GGAGAACTGA
                                                                              -4671

GTGAAAAAAA GAAATACATT AACTAGATTG GATGCAAAGT GCCTGCTGGT CATGGGTGTT TTCTGCTGGC
                                                                              -4531

CCCTGTTCAT CTGTGCCTGT TAGCCCACCC ATGGGTGAGT GGGGCAAAGT GGCCAAACTG ATTCTTAAGA

N479

GAGGCATACA TGCAGAATCC AAGTTAGTCA TGATTTCGTT TCTAGTCTGA GTGAATGTGT GTCCAGAATA
                                                   -4461
-4494

TTTTATAAAC TTTATCAGCT CAGAGGGGAA AACCTGTCTC CATACTACGT GGTTTATACA AAGCTGTCAG
                                                                              -4391

GAATTCAGCA TGATGAAGAA ATGCACAAAA CAAGTGTGAA CAGATAAGTA AAAGGATCTA CTGAAAATCT
                                                                              -4321

TCAGGGTAGT ATATTGTGTG ACAGGACCAA GAATTTGAAG TCAACATCTG TATTTGTGCC CTCTGGACAA
                                                                              -4251

AGGTATTATC CCTGATGATA TAAAAATTAA TTTTGGGCTG GGTGTGGTGG CTCATGCCTG TAATCCCAGC
                                                                              -4181

ACTTTGGGAG GCTGAGGAGG GTGAATCGAC TGACGTCAGG AGTTGGAGGC CAGCCTGTAT CGACTAATAA
                                                                              -4111

TACAAAAAAA TTAGCTGGAC ATGGTGGCGT GCACCTGTAA TCCCAGCTAC TCAGGAGGCT GAGGTGGGAG
                                                                              -4041

AATTGCTTGA ACTCGGGAGG CCGAGGTTGC AGTGAGTCGC ACCACTGCAC TCTAGCCTGG GCGACAGAGT
                                                                              -3971

GAGACWCCGT CTCAAAATAA ACAAAATTAA TTTCGAGGCC AGGTGCAGTG GCTCCAGGTG CGGTGGTTCA
                                                                          N788
                                                                              -3831

TACCTGTAAT TCCAGTGCTT TAGGAGGCCA GAGGATTGCT TGAACCCAAC AGTTCGAGAT CAACCTGGGC
                                                                           -383

AACATCAGTG AGACTCCATC TGTAGAAAAC AATCAAACAG ACAAACAACA ACAACAAAAA AACCAGAGGT
                                                                              -3761

GGGAGGATCA CTTGAGCCCA GGAGTCCGAG GCTGCAGTGA GCTATGGTCA CGCCACTGCC CTCTAGCTTG
                                                                              -3691

GGCAACAGTG CCAGACTCTG TCCTTAACAA CAACAACAAC AAAAATTAAT TCTACTTTAA CTGTCAGTTT
                                                                              -3621

CATGATATCC TTCTATTAAG AAAAACCTTT TCTATCTGAT GAACTATTGG CTAGGTTTTC TTTCTCTCTG
                                                                              -3551

CTTTTGACTA ATGCATTTAA TTACTTTCAT TTGCAAACTC TATCCTTCTC ATCAACTTTG TATTTTAGAT
                                                                              -3481

                                                                              -3411

GTGTCTATTG ACAGCCTGGC TTCCCTCAGC GATCATTATG ATGATCAAAG TAGATGAATA GGTAAAATTC

-3341

AATGCAAATA TTCCAGGGCA TCTAAATCCA TACCCCAAAT GGGAAAAGGG GAGAATTGGA AGCCAGCAAT

-3271

TTGAACACAT TACTATGGAT GTATTTTTCT CATGCGGGGG AAAAAGTGAT TTGGAGAGAG AGAATTATGA

-3201

ATGCATGTGA AGAATAAAGC CAAATTTCCT GGGAGGAGGG GAAGACCAGG AGAAACAAAA CCAAATCCTG

-3131

GCTGTGGCCT CTAAGGCATG GGGACCTGGA GTTATGCTCT CCAGGCAGAC ACAGCTCATT CTGGAGAAAG

-3061

GCTGCAAAAA TATTCTCCTT CACATTGATT TGAAAACAAT TATTAAATTC TTGTTTTCTT ATTTATCTAA

- 30l3 N487
-2991

GTGTAACTTT TTAAAACTTA CTGAGAGAAG ACGGGCACGG TGGCTCACTG CTGTAATCCA GTACTTTGGG

-2921

AAGTCAAGGC AGGTGGATCA CCTATGGTCA GGAGTTCGAG ACCAGCCTGG CCAATATGGC AAAACCCCGT

-2851

CTCTACTAAA AATACAAAAA TTATCAGGTG TGGTGGTGTG TGCCTGTAAT CCCAGCTACT CGGGAGGCTC

-2781

AGACAGGAGA ATCACTTGAA CCTGGGAGGC AGAGGTTGCA ATGAGCTGAG ATTGCACCAC TGCACTCCAG

-2711

CCTGGGCGAC AGAGCAGGAC TCCATCTCAA AATAAAAATT AATTGATTAA TTAATTAAAA ATTTACTGAG

-2641

AGCTGGTGGT TCCTTTAAGG GTGGAGCCGC CATCAAGTCC CCAGAGGATG CCCTGAATTT GGGGGCATCA

-2571

CCTTCAGCTG CTGTGGACTC TGAGCCTTGG CAGCTCCAGC TCCAGGCCTG GGAGAAAGAT GATTTCCTGG

-2501

CAGCGTGCAG TGATTGTGAG CATTTGACTA CCTTACTGCA TTTTGCCCTT ATCAKTGCTC TCCAAACATG

-2431

AGTGGAAAAC AAAAAATTTT GCTGAGACAA GCGATAATAC GAGTTAGGGA AAGTTGGAGA ATTTTATAGT

-2361

TGCTGATATC AGCAAATCGT GAGTTTCAAG CACTAACTTA CAGAAGGAAG TCCAAAATTA AAGGGGATAT

-2291

AGAAATGTGT AAAAGATGAG GTGTGGTGAA GATGGAGAAA ATGAAGAGCT CTTTAAATTT CTGAATTATG

-2221

AAGAATCACC AACAAATTAT TTTGTGGTTC CAAATACAGG GAGAAGTTCA CAGATCCACA GAACTGATGA

-2151

CAGGGTGCGG CCAGCCACAA ACCTTTCAGC ACAAGAGGGA GAAGGCTGCC GCTCCACTTT GCCTGGGCAG

-2081

TCTTTGTAAG GCAGTAGATA AGTCAGCCTC GAAGTTAGCA ATCACAGCCC TCGGCTCGGT TTCCTGCAAG

-2011
GGCATCGTTA ATGCATCACA ATTAATTTCT TCTGTCCATT AAATGTCAGC TCTCAAGTAA ATTGATGTAA

-1941
AATTTTTGTA TAGAAAACTA TTTCATATTA TTTGCACTTG ATGTTTAATT ACATTTTAAA TGTTTTGTTT

-1871
GTTTCATTTT GTTTTGTTTT TGAGACAGAG TCTTGCTCTG TTGCCCACGC TGGAATGCAG TGGTGTGATC

-1801
TTGACTCACT GCAACCTCTG CCTCCTGGGT TTAAGCGATT CTCCTGCCTC AGCTTCCTGA GTAGCTGGGA

-1731
TTACAGGCGT GCACCACCAT GCCTGGCTAA TCTTTGTATT TTTAGTAGAG ATGGGGTTTC ACCATGTTGG

-1661
CCAGGCTGGT CCCGAACTCC TGACCTCAAG CTATACACYT GCCTCAGCCT CCCAAAGTGC TGGAATTACA

-1591
GACATAAGCC ACTGTGCCCA GCCAAATGTT TTAAATAATT GTCACATATA TATACAAAAT AATTTATGTT

-1521
ATAGGTAGGG ATCTTGTTAT ATTTTAACCT TCAAAGTATA TTCCTAAGCT TTTTATTTAT TTTTTATTTT

-1451
TTATTTATTG AGACAGTCTT GCTCTGTCGC CCAGGCTGGA GTGCAGTGGC GCAATCTCGA CTCACTGCAA

-1381
ACTCTACCTC CTGGGTTCAA GCGATTCTCC TGCCTCAGCC TCCTGAGTAG CTGGGATTAC AGGTGCGCAC

-1311
CACCATGCCC AGCTAATTTT TGTATTTTTA GTAGAGACGG GGTTTCACCA TATTGGCCAG AGCTGGTCTC

-1241
AAACTCCTGA CCTCAGGTGA TCCATCCACC TCAGCCTCTC AAAGTGCTGG GATTATAGGT GTGAGCCACT

-1171
GCGCCTGGCC TATTCCTAGC CTTTTATATA TAGACCTTTT TCTTTTTCAC ATTTTAAAGG AACTTTATG

-1101
TTTAATCATG GAATATTTCA AACATACAGA AAAATCACAG AAAATAAATA ACAACCACTC ATTTATCTTC

-1031
TCCCCAACCC CATGTAATAA ATATTAAAAT ATTGTGTTAA ATGCTAAATT TAACACATGC TAAAGGTTCC

-961
TGGCTGGATG TGGTGGCTCA CGCCTGTAAT CCCAGTACTT TGGGAGGAGG AGGTGGGAGG ATTGCTTGAG

-891
TCCAGGAGCT CGAGACCAGC ATGGGCAACA TAGTGCGATC TCGTCTCTAC AAAAAACAAA AAAATTAGCT

-821
GGGCATGGTG GTGTGCATCA GTAATCCCAG TGACTGGGAG GCTGAGGTGG GAGAATTGCT TGAGTCTGGG

-751

AATTTGAGGC TGCAGTGAGC CCTGATCATG CCACTGCATT CCAGCATGGG CGACATAGCA AAACTTGTCA

                                                                        -681
AAAAAAAAAA AAGTTTCCTC TCTGCCCCAC CATAGACAAC CACTCTTCTG ATTTCTATCT TCGTAGATGA

                                      AP1   -611
ATTTTGCCCA TTCTCTTGTA TATGAAAGGA ACCAGACATT AGGCATTCTG GTGTCTGGTT TCTTTCACTT

                                                                        -541
AAGATAAAAT TGAGTTAACC TGTATTGTTG TACAGAACTG CAGTTTGTTC TTTGTTATTT ATTGTAAAGA

                                                                        -471
CAGGGTCTGG CTATGTTGCC TAGGCTGGTC TCGAACTGTT GGCCTCAAGC AATCCACCTG CCAAGCTCTG

                                                                        -401
GGACCACAGG CATGAGCCAT GGCATCTGAT CKGTAGTTTG ATCTTATTTC TTGCTGAGTA GTAGCCCATG

        AP1   -331
GCATGACTTT ATTATTTTGG GTGTCCATTC TCCTCTGGAG GGGCTCTGCT TTTTGAAACC ACACCCTGGC

              Cts   -261
CTAGCTCCCC TTCTCCCTGC CTCTCTGCAG GCTCACATCC ACATGCCAAG ACCTCTGCAG CCATTCTGCT

        Cts   -191
TCCTGTCCTT CCACTCCTGT GGGACCTCAG AGAGCTACGG GGCTCCCTGG GTACCAACTG GCTCCTGAGG

        Sp1/Sp3   Sp1/Sp3   -121
CCTGGGGGAG GGTGGTCTTC TGGGAGAAGG AAGCCAGGTC CCTGCAGGTT GTGGAGGGGG ACAGAATGAG

                         Sp1/Sp3   Cts -51
GGTTTTTCCC CAGGATGTTG TTGGCCCCTG CCCCCACTTC TGTTCCATAA TTAACCACGC CCCTCCTACC

                     Sp1/Sp3   +1   20
CACTGTGCCC CTCTTCCTGC TGTGTGGAGG CCCTGAATCA TTATTTTAAC TACCCCCTGG GAGGGTGAGC

      Cts   Cts   90
ACCTTCTGTG CTCTGTCCCC AACCTTCCAC TTCCCCTCAA CGCGCTGCTC AGGGATGACC TTGGGCACTG
                                                        M   T   F   G   T

                                                                      160
TGCTTCTTCT GAGTGGTAAG TGGGGCCAGG GTGCTGGGGA GAAGCTTGGA GGAGTTCTGA GGGGACTCCA
V   L   L   L   S   V

                                                                      230
TCTGGGAGGG CAGGCTGGGG GCTGGTGGTC GGCTCCAACC ACTCTTATGA GGAGCTGAGG CAGGGGAGTG

                                                                      300
CTTCATGTGC GAGTGGCCCG GAGTCAGTAG AGTGTGACCT GAATGAAGAG GGGCTCAGGG GCTGTGCTCA

                                                                      370
GGTGGCGACT AAGCTACCTC TCCAGCTGGC TATGTTGTCC CAGGCTTCCC TGCTCCCACT CATGGAGTCC

                                                                      440
CTGGTGTGGG TGACAGAGGT CTCCCCAGCC TCCCCCGGGA GTGGAAGGCC ACAGAAGCCA CCAGGGAGGG

                                                                      510
GGAAAGGTTG GACATCACCT CCCTGGGCCT NNNNNTTCCC CCAAGTCCTG ACTGCACGTA GGGAAGAGGC

                                                                      580

CCCCTGCTGA AAACTGCATC AGAGTCACAT TCACGTGCCA TCAAAAATCA GGCTTGGCTG GGTGCGGTGG

650

CTCATGCTTA TAATCCCAGC ACTTTGGGAG GCCGAGATGG GCGTATCCCC TGAGGTCAGG AGTTTGTGAC

720

CAGCCTGGCC AACATGGTGA AACCCCATCT TTACCAAAAA TATAAAAATT AGCCGGGCAT GGTGGCGTGC

790

ACTTGTAATC CCAGCTACTT GGGAAGCTGA GGCAAGAGAA TCGCTTGAAC CCAGGAGACG GAAGTTGCAG

860

TGAGCTGAGA TCGTGCCGTT GCACTCCAGC CTCAGCAACA GAGCGAGACT CCATCTCAAA AAAAAAAAAA

930

AAAAAAGAA AAAAAGAAA AAGAGGCTGG GAGGTCCTAG GGATTGGGGC TTCTTTAACT CCCAGCCTCC

1000

CCGCCCACCA AATATTCCTC AGTCTGGCT TCTTATCATG GATTCAACCT GGATGTGGAG GACCTACGA

                 L A S Y H G F N L   D V E E P T

1070

TCTTCCAGGA CCATCCAGGC GGCTTTGGCC AGAGCGTGGT GCAGTTCGGT GGATCTCGGT AGGCCCCACT _Intron 2

I F Q E D A G G F G Q S V V Q F G G S R (3019)

1140

CACCCTCCTT CCCCAACCTC CACTACATCA AGTCCTGTGG ATGGGTACAC GTGGGTTACC CGAGGGAGGT

1210

GTCCTGGAGG AAGGCCAGCA GGGGTGAGAA GTCTTCCCTT GGCTCCTTGG AGGCCCTGAC ATCAGCACCT

1280

ATTATTCTCA ATCCCAGGAA AGGCCACAAA ACTCTAGACA AGACCCTACC TTACCTCGGG AGGGAAGCCT

1350

TGAACCTGCC TCCCAGGCAG GGCCCACTTC TTGGGGCCAG TATGGTCACA CAGGGCCCAC ACTCATTAAC

1420

TTTGGAGTTT AATGTTCTGC CCTTGACCTC TTGAAATTCC TGATTATTTT TATTTTTATT TTTACTCCAG

1490

CTCTGTTACC CAGGCTGGAG TGCAGTGGTG CAATCACAGC TTACTGCAGC CTCAAACTCT CGGGCACAAG

1560

TGATCCTCTC ACCTCAGCCT CCTGAATAGC TGGGACCACA GGTGCATGCC ATCATGCCTG TTTTTTGTTT

1630

TGTTTTGTTT TACTTTTTAC AGAGATGGAG TCTTGCTATG TTGTCCAGAC TGGCTGAACT CCTGGGCTCA

1700

AGCAATCCTC CTGCCTTGGC CTCCCAAAGT GCTGGGATTA CAGGTGTGAG CCACCCTGTC TTGCCAATTC

1770

TTAAAAATTT TATCTGTGCA TTTGTGTTTT GCAAGTAAAG AATGATGGCA GGGCTGGGCA CCATGGCTCA

1840

CGCCTATAAT CCCAACSCTT TGGGAGGCTG AGGCGGGCAG ATCATCTGAG CCAGGAGTT TGAGACCAGT

1910

TTGGCCAACA CAGCAAAACC CCATCTCTAC TAAAAATGCA AAAAAATTA GCCGGGCATG GTGGCAGGCA

1980

TCTGTAATCC CAGCTACTTG GGAGGCTGAG GCAGGAGAAT CGCTTGAACC TGGGAGGTGG AGGTTGCAGT

2050

GAGCCGAGAT CGTGCCACTT TACTCCAGCC TAGGTGACAG AGTGAGACTC CGTCAAAAAA AAAAAGTCAT

2120

GGGAGAAGGG AGATGTACTG GGGGTTTGGA GCCTTAGCTC AGCAGCAGCC CCACCTCCCA CCGCCTCCTG

2190

AAGGGTGGTG AAGGGGTATC AGCTGCTGGC TCCCCCACCC ATGTGGGAGC AATGACCGCT GCTACCTTCC

2260

GCCCCTGGCA TGAGCTGGGT AAAGTCAGTT AGGGGCGCTC ACTCTGGGAG TACCCCGAGG GAGTGGGACA

2330

CTACATAGCA AATAAAAAAC GTCAGGACAG GTTGAGGAAA GAGAGCAGAA GAAAGGTAAG AGCCCCCCAA

2400

CCCCAAGAGA CCCCACAGTT TTATTTCAAA TTGGGACCCA CAAATTATGA ACCTGCCCCC ACTTCCAGGA

2470

GCTCACATTC TCCTGTCCCA GAGAGTTCAA GTCACAATGT GACACAGGTG TCACCAAGGT CTGGGGGGCG

2540

CAGGCAGGGA GAGAGCAGAC CCAGGAGGGT TCCATGGAGG AAGTGGTGCT GGCAGTGAGC CCCAGTGGAC

2610

AGGAAGGCTC AGTTGGTCAC GAGGAGCTAT AAGAGGTCAC CGAGCTCCAA CCGCGCACCC CTCTCCCTTC

2680

CTCATGTGAC TGGCAGTCTG GGGGGATGGA AGCAAGCACC AGGCACCAGG CTTTTGTTTT TCTTTATTTG

2750

GAAATGTGGT CAACTGAGGT GCACAAATCT GAAAGACCCA ATCTGATAAA GGATACACAT GTGCGTGCCT

2820

GGGTGAGCCC CACCTAGGTC AGCTGCTCCA GTGTCAAATC CCACAGGCAC AGGGCTGCCG TGGACCCCTT

2890

CTCATCACCC AACATCCCCA GAGAACCCCT GGTCAGACTT CTGTCACCAT CAGTTTTTTG GGCCACATTT

2960

TAAAAAAAGA ATACATTGGC TGAGTGCAGT GGCTTATGCC TATAATCCTA GAACTTTGGG AGGCTGAGGC

3030

GGGTGGATCA CCTAAGGTCA AGAGTTCAAG ACCAGCCTGA CCAATATGGT GAAACCCTGT CTCTACTAAA

3100

AAATACAAAA ATTAGCCTGG CGTGATGGCA GGTGCCTGTA ATCCCAGCTA GCTGGGTGAC TGAAATAGGA

3170

GATTTGCTTG AACCTGGGAG GTGGAGGTTG CAGTGAGCTG AGATCACGCC ATTGCACTCC AACCTGGGTG

3240

ACAGAGTGAA ACTCTGTCTC AAAAAACATA TGGGTTGATG GGTTACACTA AAGTTTTGCT CATCGTTTGT

ATCAGCAGGT TCCAAACTGC TACCTCTCTA GCCAATGCTC AGATTTTCTT CACAAAGCCT TAGGCATCCC
3310

CTGAATCATG ATGCACAGGG ATTGTAGCTT TCTGTAAAGG AGCGGCACCT AGAAGGAACC CTCACATGGC
3380

CATTTAATGA AGCCTTGCTT GGCGCATTAA AATACACCAG TATCTGTCTG CTTTTCTCAC AGACAGGAGA
3450

TTGTGGGTAG TGAGAAAACA TTTCCAAAAT TAAAAAACTT TCCCACTCAG GGAGTTTTGC AAATAAACCC
3520

TTGACTCTAC ATAACTATAG ATATAGTTAT GGATCCTAGT ACACTGCTTT ACATTGGCCA ATTGAAATTG
3590

CTTATACAAT ATTTAAATTG GTCCAATGAA TTACAGAATC AACTATTTGT TTTGAAAGCA CATGTCTTCA
3660

GGAAATTGTT CCAATTAACT TGACATGATC TTATTCTTG GGTGGTTCAA AATAATGGCA ACTCAGAAAC
3730

GCAATGTGCT TACCCATGAT TGGGAAATGC CATTTTGGTC TTTAAATAGG TCTTTTTTTT TTTTTTTTT
3800

TTTTTTTTT GGTGAATGTT AAAAAGAAAT TTCTAAACAT AAATACACAC ATACGTACTT ATGCACACTC
3870

AAAACCAAAT AAACCCCAGC ATGGCCCCTG GGCATCTGTG AGTTACACTT GGGCCCTGAT TTCTGAATAT
3940

TCTGCCAAGT GGCAAATGCC AGGAATTTCC CCCACAGAGT CTCGCTTCCC CATGGAGGGA CACTTCCTCA
4010

CCCCCAAGTG CCCGCTGCTC CCACCCCTCC TGTGGCTGCA GTGACATGGC CATGGTTGTG TCTCCAGACT (···)
                                                                                L
4080

CCTGGTGGGA GCACCCCTGG AGGTGGTGGC GGCCAACCAG ACGGGACGGC TGTATGACTG CGCAGCTGCC
 V V G A P L E V V A A N Q T G R L Y D C A A A
4150

ACCGGCATGT GCCAGCCCAT CCCGCTGCAC AGTGAGTGAC CACCTGGGAA TTGGGCCCCT CAACCCTCCT
 T G M C Q P I P L H I
                         Intron 3
4220

GGACCCAACT GTGCCCCCGC TTAGCTTCCA GTCCAGACCT TCCCCGCAAA TGAGTGTGTG CTGTGAGTGA
4290

GACCCCGCGT GTCTGCCCTT GCAGTCCGCC CTGAGCCCGT GAACATGTCC TTGGGCCTGA CCCTGGCAGC
                        R P E A V N M S L G L T L A A
4360

CTCCACCAAC GGCTCCCGGC TCCTGGTGAG TGAGTGTCTT GGGCCACGGG GGGGTGGGGT GGGGCGGGGG
 S T N G S R L L
4430

GTGTTGTTGG GGAGGAGGCT GGGGCTGGGA GTGAAGGAGG AGGGGCTGCT AGGGACTCCT GGCTCACAGG
                        Intron 4
4500

4570

```
CTTCTGCCTC CAGGCCTGTG GCCCGACCCT GCACAGAGTC TGTGGGGAGA ACTCATAGTC AAAGGGTTCC|
            A   C  G  P  T  L    H  R  V    C  G  E  N    S  Y  S    K  G  S
                                                                              4640 (1700)
TGCCTCCTGC TGGGCTCGCG CTGGCAGATC|ATCCAGACAG TCCCCGACGC CACGCCAGGT AGGTCCCTGG  Intron 5
  C  L  L    L  G  S  R  W  E  I   I  Q  T    V  P  D  A   T  P  E
                                                                              4710 (4267)
CAGGCCATGG TTCCCTGTGG AGCACATGCT GGCACTGAGG GTGAGCAGGC GTGAGGCCTG TGTCTGGGCC

                                                                              4780
CCTGTGCCCT CCCTGGAGGG CCGAGTGTGG CTAGGAGAGA AGCCAGGAGA AGAGGGTGGC TCAGGCAGGA

                                                                              4850
GCCCTGCTGC TCCAGGGTAG AAGTTCTTTG CAGGGTTTTT CTTTATATTT TTTTCTTTTT AAGACAGGGT

                                                                              4920
CCCTGCCAGG CACAGTGGCT CAGGCCTGTA ATTCCAGCAT TTTAGGAGGC TGAGGTGGGC GGGATCACCT

                                                                              4990
GAGGTCAGGA GTTCGAGACC AGCCTGGCCA ATGTGGTGAA ACCCCTCTAC TAAAAATACA AAACAAAACA

                                                                              5060
AAACAAAATA GCAGGATGTG GTGGTGTGCG CCTGTAATCC CAGCCACTCG GGTAGGCAGA GACAGAAGAA

                                                                              5130
TCGCTTGAAC CCAGGAGGCG GAGGTTGCAG TGAGCTGAGA TTGTGCCATT GCACTCCAGC CTGGGTGACA

                                                                              5200
AGAGCAAAAC TCCATCTCAA AAAAAAAAAA AAAACAAAAA ACAGAGTTTC TGTCAGGCTG CATGCACCAC

                                                                              5270
CACACCCTGC TAATTTTTTT GAGACAGAGT CTTGCTCTGT CGCCCAGGCT GGAGTGCAGT GGTGCAATCA

                                                                              5340
TAGCTCACTG CAGCCTCGAA CTCCTGGGCT CAAGTGATCC TCCTCCCTTA GCCTACTGAG TAGTTGGGAC

                                                                              5410
TGCAGGTACA TGCATCACAC CTGGCTAATT AAAAAAAATG TTTTTGTAGA AATGGGGGTC TTGCTATGTT

                                                                              5480
ACCCAGCCTG GTCTTGAACT CCTGGGCTCA AGTAATCCTC TGCCACAGCC TCTCAAAGTG TTGGGATGAC

                                                                              5550
AGGCATGAGT CCTTGTGCCT GGCCTGAGGG ATGAAAGTTC TGATGGAGGC AGAGAGGAGC CCCACTGTGC

                                                                              5620
GGGCTGTAGA GGGCACAGCA TCTTCCAGTT GCCAACAGGT GCATGGCCAC TTCTTGAGTT TCAGAGGAAG

                                                                              5690
GACCTTAGTG TGGTAAAGAA CGTGGTGAGG AAGATAAATC CATGAGGGAG GTGTTTCTTC TGGATGGTTC

                                                                              5760
ACTGCTGAGC TTCCAGGATT CCCCAAACTA ACTTTCCTCT CGAAGAGGAG CAAATGACAG GGCTGCGGAA

                                                                              5830
AATGCGATGT GCAATTTTGT CAGTGCCCAT GTCTTCCACA GAGAACAGGG CCTGGGGACA CCACCATGAC

                                                                              5900
```

ATCTCTCTGA GGGTTGGTCT GCATCATGGT GGTTCCCAAG TTTGTTTTCC ATGGGCACCA GGCTTCATTC

5970

CCTTGAAGCT TCATTCCCTC AAAGCCATTC AGTTTCCTCA TTGGTAAAAT AGAGCTCAAT AATCAGGGGG

6040

TTATGAAGGT GAAAGGGATT GAGGTGCATA AAGCACTTGG AACCCTGCCT GGCACATAGT ATGTGATAGC

6110

CCCTCTGACC CATCTTCCAG CTGGGGACTG CATGCTGGGA CTGGGAGGAA GATACAGGCA AACTGTCTCA

6180

TCTGCCGTGT GAGAGGGAAT GCCAGGGGCC GCTCAGGGTG CTGACCGAGG GTGGGGCTTC AGACCAGAGA

6250

GGCCATGATG ACAGGCATGC TGGGCCTTTA GACAAAGGTG GAGCAGCAGC AGAAACATTA CCAGAGCAAA

6320

TGGTGAGGGT GGAGTCTATG GAGGGGACCA AGGGAAGGGG GAAGGGACAT CCAGGGTTCT TGGGGGGACC

6390

GTGCCCAGCC TGAGATGTCT GTGAAGCTAG GTTAGGGAGG TGGCACTTAA AAACAAGGGG TAAATGTCTT

6460

CTCACAGCCA TCCGTGGAAC TCATGAGGTG GGATGCCTGA TGCAAATGGG ACTGGAGCAC AAAACTGGTG

6530

CAGGCAAGGG GGGTGTGGGT CCAAGTAGAA GGGACCAGGG TCCACTGAGG ATCACCTGTG TGCCAAGCAG

6600

TGCTGAATAC CTGGTATGAA TCACCTTATT GCATCCTCAC AACATCCTGG GTGGTGGGCA GGCCCATTCT

6670

CATTTTACAG ATATGAAAAC CAAGGTTCAG ATAGATGAGT TCCATCGATA GCAAGAGGCA GAGCCCAGAG

6740

CTTGAGCCAT CCTTGCCTGA TTGGTGGGGT CCTTTTTCAA AAGGATAAGT CCAGGCTTCT GCTAGTGGGA

6810

GACCAGGGGA TACAATAAAA AGACCAAGAA ACAGAAGAGA CATTGTGAGA GGATTTGCCA CAGACCTGGC

6880

CTGAGAGAGG ATGAGAGGGT GGTTTCTTGA CGCAGCTGAA AAAACAGGCA CCACTGCAAG ATGTTGGCTG

6950

CCCAGATGTG GGCAAAAAAC GGGGAGCTCC TGGGGGGATC TGCAGCCTGC CCCATGGATG TCAAGATTTG

7020

CTGGTGATTG AAGAAGCAGG AAGGAAGTGA CCTTCTGTTT CTCCCCAGCA CCCTTGAAGC ACCAGTGGTT

7090

GAGCAAGTGG GGTAGGGGAG AGGAAAGAGG AAAAGGCATT TTTTTTTTCT GCAGTGGTGG GCAGGGGGCA

7160

GAAACCACAG CCCTGTGGTG TGGGCCTCAC ACCTTAGTGC TCTGGTGGCC TGATCTCCCA GTGCCCTGCG

7230

GGCAGCACAG GATGTGGCTG CTGGTGGAGG TACCAACTGG GCCCTGAACA CAGGCCACAC ACCCCCCATG

7300

AGCCTGGGGA CAGCATGAAA AGTCTTATTT GTTCATGTGC ATATGATGTG CCCTCACGAT TGCAGAGTGA

7370

ACTCCACAAA CTCTGAGGTC ACTTGGGAAT GTTCTTTTTT TTTGAGACGG AGTCTCACTC TGTCGCCCAG

7440

GCTGGAGTGC AGTGGCACAA TCTTGGCTCA CTGCAGCCTC CACCTCCCAG GTTCAAGTGA TTCTCCTGCC

7510

TCAGCCCCCC AAATAGCTAG GATTACAGGC ACCGCCACCA TGCCGGGCTA ATTTTTTTGT ATTTTTAGTA

7580

GAGATGGGGT TTCACCATGT TGGCCAGGCT GGTTTTGAAC TCCTGACCTC AAGTAATCCG CCCACCTCAG

7650

CGTCCCAAAG TGCTAGGATT ACAGGCGTGA GCCACCACTC CCAACTGGGA ATATTCTTGG GCACCGCACC

7720

CATGGGAGCA TGAAGGGTGG ATGCAATGCA ATCATAACAG AGGCCCAAGG TCAGCACTGG GGTGCTTGCC

7790

TGTCATCCCA GTGCTTTGGG AGGCCGAGGT GAGTGGATCG TTAGAGCCCA GGAGGTTGAG ACCAGCCTGG

7860

GCAACATGGC GAAACTCCGT CTCTACAAAA AGATACAAAA ATTAGCCAGG CAAGGTGGTG CACACCTGTA

7930

GTCCCAGCTA CTCAGGAGAC TGAGGTGGGA GAATTGCCTG AGCCTGGGGA GGTCGAGGCT GCACTGACCT

8000

GTGATCACAC CACCACACTC CAGCCTGGGT GACAGTGAGA CTCTGCCTCA AAAAAACAAA AAATGAAAAA

8070

ACCAGAGGCC TCAGCCAATG CCTGGGGGGC TCAGAGTGCA GCTGGCCCTT CAGACGCTGA ACCAGTCATC

8140

GGTAAAGGTT TCCTCCAGGG GCAGGAGGTG TCCCAGTGGG CAACAGTTCC CCTCTGCCTA GCGTGATTCC

8210

TGGGAAGGGA CTCAGCTCAG AGCCAACTCC ACGTAGCTGG AAATAAGGAC CTCTGACCGA CTGGGGGTAG

8280

GGTGGGGTCT GGGGTGGATC CCTGCCCCAC CCCCACAGCA TCCCTACAGG CATATCCTAC AGGCCTCGAA

8350

GGTGCCTGGC ACGTGGTGAG AATGGTGCCA GCGGCTGACC CTGGCAGAGG GCCAGGACTT GTCTCCAGCA

8420

CCCATGTGCG TGTTGCTTTA TCCTTGCAGT GATCCCACGT GGTAGCCACT GATATTACCT TCATTTTACA

8490

GATAGGGACA CTGAAGTCCA GAGAAGTTAA GTAATGTGCC TGAATTCACC AATTAGCACG TGGCTGAGCT

8560

GGGGTTTAGC CCGGCACAC TGGCTCCAGA ACACGCGCCC TGAACCACTT TGCTAAACAT TCGCCCTTGA

8630
TGCTTGTGGC ACCCCTAGCA TCTGTGTTTA ATGAATATTT GTTGATTAAA TGGATGAGGA GCCCACCTGG

8700
GTCCTGTGTT GTCATCCCTC TCTTTCCAGG CCATGTGGGA GGAAGGGAGC AGGGGGCTGG GGTGGCAGAC

8770
TGGGGCCTCC TCCAAGGAGG GGTCGGAAAC TAGGTGGGGA TGCCAAGAAC AGCCCCCGGG CTCTGTTGAG

8840 (1770)
CAGGAGCTGC AGGAGGGGGT TGGGCCCCCG CAGTGCATCT CCGATTCCTC CCCATTCCCC CACAGAGTGT
C

8910
CCACATCAAG AGATGGACAT CGTCTTCCTG ATTGACGGCT CTGGAAGCAT TGACCAAAAT GACTTTAACC
P H Q E M D I V F L I D G S G S I D Q N D F N Q

8980 (1910)
AGATGAAGGG CTTTGTCCAA GCTGTCATGG CCCAGTTTGA GGGCACTGAC ACCCTGGTGA AGACTGGGCA
M K G F V Q A V M G Q F E G T D T L

9050   Intron 1 (1255 b)
CCTGGGGCTG GGGTTTGGGG GACGGGGGAG CTGGCCTCG GGGAGGCATC CCGGGAGGGG TGGGGGCAGG

9120
CCAGTGAGCC GTGTGTGATG GGGCTGGGGT GGAGAATGAA GCTATGGTCC CAGCACAGGC CCAACTTGAG

9190
CCCTGACCTA GGAGGCCCTC TTGGCATTTA ATAATGTATC CAAAAGCTAC AGGAAATACA ATGTTTCTGC

9260
TGTTTAAAAA CATTGAAAGT GTTTAAAAAA TAATTTGCT ATGGAATTTG ACTGACATAT TTTCTATGTA

9330
ACTAATGATT TCTTTTTATT ATTTTATTTT CCCATAACTA GTGCCTATGT ATGTGCAAAA AAACCCCACT

9400
GTTTTGTTTT TTAAACAGGG TCTTGTTCTG TCATCCAGGC TGCAGTGTGG GGGTGCAATC ATAGCTCACT

9470
GCAGACTTGA TCTCCCAGGC TCAAGTGATC CTCTTTCCTC AGCCTCCTGA GTAGATGGGA CTACAGGTGC

9540
ATGCCACCAT GCCTGGCTAA TTTTTAAAAA TTTTTATTGT AGAGATGGGG TCTTGTTATA TTGCCCAACC

9610
TGGTCTCCAA CTCCTGGCTC AAACAACCCT CCCACCTTGG CCTCTCAAAG TGCTGGGATT ACAGGCATGA

9680
GTCACTACGC CCAGCCACAA ATTCTTGAGA AGGGAGGAAA TACACAAATT GATTTAATGT GTGATAAAAT

9750
TTTTAAAAAG CTAAGGGTAG CAAATATTGT GGTTGGCAAA GTTATTTTTA ATAAACTGTG AAGGAAATAT

9820
CATTTGGTTT AGATTTTCCA GTTTATATTT TGTATTTTTG GGCCAATGAA TTTTCTTCTC TGGGTTTCAG

9890

TATATTATTT GTAGGGCTTC AAAACACATG GGAAATGGTT TGTAAATCCA AAATAATTCC AAAATAAAGT

TTATTAAAAC TGAAAACAAT ATGGCTTGGT GTGGTGGCTC ACACCTGTAA TCCCAGCACT TTGGGAGGCT
9960

NGAGGTGGGA GTATTGCTTG AGGCCAAGAG TTCGAGACCA GCCTGGGCAA CATAGTGAGA CCTTGTCTCT
10030

ACCAAAAACA AAACAAAACA AAAACAAAG CCAGGCATGT GACGTGTGCC TGTAGTTCCA GCTACTTGGA
10100

GGCTGAGGCA GGAGGATCAC TTGAGGCCAG GAGTTTGAGA GACCCTGTCT CTACAAAAAA TTAAAATAAA
10170

AACAATAGTA ACAGGCACTG AGCCCTGGGC CCTCCCCACT GGCCTTTGCA GTTGCACTG ATGCAGTACT
10240 (142)
                                                              F   A L M Q Y
CAAACTCCT GAAGATCCAC TTCACCTTCA CCCAATTCCG GACCAGCCCG AGCCAGCAGA GCTGGTGGA
10310
S N L L K I H F T F T Q F R T S P S Q Q S L V D
TCCCATCGTC CAACTGAAAG GCCTGACGTT CACGGCCACG GGATCCTCA CAGTGGTGTA AAGCAACCCC
10380
P I V Q L K G L T F A H A E I A T V
GACCCCA... Intron 7   T   G   L